Structure, Vol. 12, 893–903, May, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.str.2004.03.015

Structure of Protein L7Ae Bound to a K-Turn Derived from an Archaeal Box H/ACA sRNA at 1.8 Å Resolution

Tomoko Hamma and Adrian R. Ferré-D'Amaré*
Division of Basic Sciences
Fred Hutchinson Cancer Research Center
1100 Fairview Avenue North
Seattle, Washington 98109

Summary

The archaeal RNA binding protein L7Ae and its eukaryotic homolog 15.5 kDa/Snu13 recognize K-turns. This structural motif is canonically comprised of two stems (one with tandem A•G base pairs, the other with Watson-Crick pairs) linked by an asymmetric internal loop. L7Ae recognizes conventional K-turns in ribosomal and box C/D RNAs but also binds specifically to some box H/ACA RNAs at terminal stem loops. These have the A•G paired stem, but lack the Watson-Crick stem. The structure of *Methanococcus jannaschii* L7Ae bound to a symmetric duplex RNA without Watson-Crick stems demonstrates how a binding site for this component of diverse ribonucleoprotein complexes can be constructed with only the A•G stem and the loop. The RNA adopts a functional conformation with the aid of a base triple and tight binding of divalent cations. Comparison with the 15.5 kDa/Snu13-RNA complex structure suggests why the eukaryotic homolog does not recognize terminal stem loop L7Ae binding sites.

Introduction

In eukaryotes, two classes of small nucleolar (sno) RNAs guide the posttranscriptional modification of cellular RNAs. Box C/D snoRNAs assemble into ribonucleoprotein (RNP) particles with 2′-O-methylase activity (Kiss-Laszlo et al., 1996; Nicoloso et al., 1996; Tycowski et al., 1996), whereas box H/ACA snoRNAs are responsible for the isomerization of specific uridine residues into pseudouridine (Ganot et al., 1997; Ni et al., 1997). In both cases, the RNP specifies the locus of the enzymatic activity by base pairing with its substrate through a segment of the snoRNA that is complementary to sequences flanking the target nucleotide (Figure 1A). Hundreds of both classes of guide snoRNAs have been described (reviewed in Bachellerie et al., 2002). Although, only few snoRNAs are essential for cell viability, some snoRNAs have been shown to be advantageous for cell growth and required for full activity of the ribosome (Badis et al., 2003; King et al., 2003). Recently, it was discovered that both types of guide RNAs are also present in archaea (Gaspin et al., 2000; Omer et al., 2000; Tang et al., 2002), suggesting that the molecular machinery of snoRNA-directed posttranscriptional modification is very ancient.

Together with the guide RNA, four proteins comprise the core of the eukaryotic box H/ACA snoRNPs. These are the pseudouridine synthase dyskerin/Cbf5 and three structural proteins: Gar1, Nop10, and Nhp2 (Bousquet-Antonelli et al., 1997; Henras et al., 1998; Watkins et al., 1998). Analysis of archaeal genomes revealed the presence of close homologs of the first three (Watanabe and Gray, 2000; Watkins et al., 1998). However, the closest homolog of Nhp2 found in archaea was ribosomal protein L7Ae (Rozhdestvensky et al., 2003) (45% identity between the human and *Methanococcus jannaschii* sequences). Eukaryotic box C/D snoRNPs are also comprised of several proteins (Galardi et al., 2002). Homologs were found in archaeal genomes for all of these except for the RNA binding protein 15.5 kDa/Snu13p. In this case too, the closest archaeal homolog that could be detected by sequence analyses was L7Ae (Kuhn et al., 2002) (41% identity between the human and *M. jannaschii* sequences). A biochemically active archaeal box C/D RNP has been reconstituted in vitro employing L7Ae (Omer et al., 2002). In addition, L7Ae has been found to bind in vitro to both box C/D and box H/ACA small RNAs (sRNAs) from archaea (Kuhn et al., 2002; Rozhdestvensky et al., 2003). Thus, it appears that in organisms of this kingdom, L7Ae is part of three different types of RNPs: the large ribosomal subunit (Klein et al., 2001), the box C/D methylases, and the box H/ACA pseudouridine synthases. Presumably through gene duplication events, an ancestral protein diverged to produce the eukaryotic paralogs Nhp2, 15.5 kDa/Snu13p, and the ribosomal proteins L7a, L30, and S12 (Koonin et al., 1994) that assemble into the corresponding RNPs (box H/ACA snoRNPs, box C/D snoRNPs, and the ribosomal subunits, respectively).

The similarity between the archaeal protein L7Ae and the eukaryotic L30 and 15.5 kDa proteins implied by sequence analyses has been borne out by structural studies. The proteins all fold into a highly conserved compact globular domain (Klein et al., 2001; Mao et al., 1999; Vidovic et al., 2000). The structural similarity of the proteins extends to their function. All three proteins bind specifically to kink-turns (K-turns). The K-turn is an RNA motif that was identified in the structure of the 15.5 kDa protein bound to its cognate site in U4 snRNA (Vidovic et al., 2000), and then found to be a recurrent element of ribosomal RNA structure (Klein et al., 2001). Its sequence signature was independently detected as part of antiterminator RNAs (Winkler et al., 2001). Analysis of occurrences of the K-turn in the 50S and 30S ribosomal subunit structures led Klein et al. to propose a consensus sequence and structure for this motif (Klein et al., 2001). The canonical K-turn comprises four characteristic elements (Figure 1B). First, the phosphodiester backbone undergoes a bend of ~120° at an asymmetric internal loop (or bulge). This results in the extrusion of a nucleotide from the loop into solvent. Second, the loop is flanked on one side by a stem consisting of Watson-Crick pairs. Third, the loop is flanked on the other side by a stem with two tandem-sheared

*Correspondence: aferre@fhcrc.org

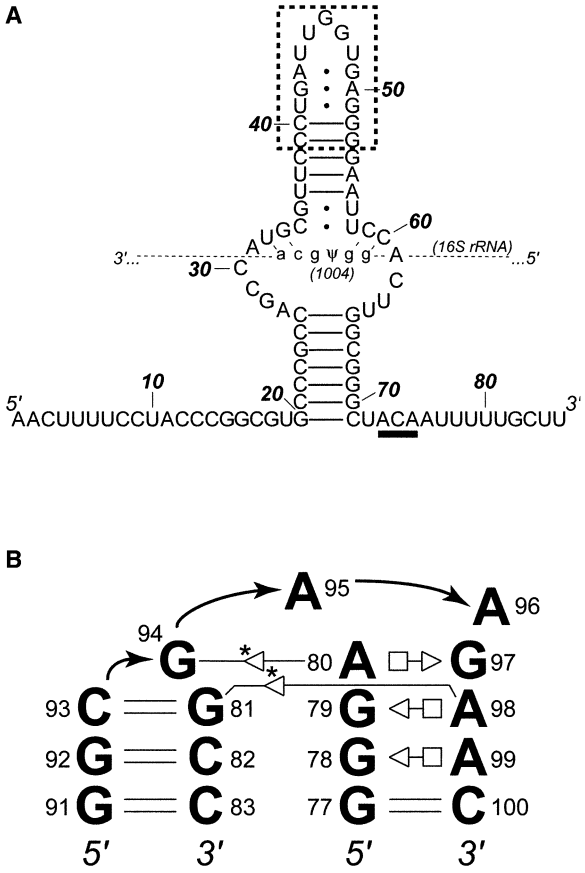


Figure 1. Secondary Structures of a Box H/ACA sRNA and a Canonical K-Turn

(A) Sequence and predicted secondary structure of the pseudouridylation guide sRNA afu-190 from the archaeon *Archaeoglobus fulgidus* (Rozhdestvensky et al., 2003). Also shown is part of the sequence of 16S rRNA that is complementary to the central bulge of the sRNA and contains the pseudouridine modification at position 1004 (Tang et al., 2002). The dashed box indicates the biochemically determined binding site of L7Ae in the sRNA (Rozhdestvensky et al., 2003) and corresponds to the 15-mer RNA oligonucleotide used in this study. The terminal K-turn motif is, in general, located five to six base pairs above the base of the upper stem (Rozhdestvensky et al., 2003).

(B) Secondary structure of a canonical K-turn, from 23S rRNA of *Haloarcula marismortui* (Klein et al., 2001), depicted using the base-pairing symbols of Leontis and Westhof (2001). The asterisks indicate two A-minor (Nissen et al., 2001) interactions between the complementary (left) and noncomplementary (or A•G) stems. Note how the L7Ae binding site (A) could form the tandem sheared A•G pairs of the noncomplementary stem, but does not have sequence that could form a Watson-Crick complementary stem.

A•G pairs. A wobble G•U pair (or another noncanonical base pair) is often found distal to the A•G pairs. Fourth, the sharp bend in the trajectory of the helix is stabilized by A-minor (Nissen et al., 2001) interactions between the two helical stems. The binding sites for L7Ae and its homologs in rRNA (Klein et al., 2001), snRNAs (Vidovic et al., 2000), and box C/D snoRNAs (Watkins et al., 2000) conform to this consensus.

Biochemical characterization of L7Ae binding to archaeal box H/ACA RNAs revealed two classes of cognate sites. Among predicted box H/ACA sRNA containing K-turn motifs identified from small nonmessenger RNA (Klein et al., 2002; Tang et al., 2002), approximately 54% of L7ae binding sites were at asymmetric internal loops that conform to the K-turn consensus (Rozhdestvensky et al., 2003). However, 46% fell in terminal stem loops (Figure 1A) with A•G pairs closing the stem (Rozhdestvensky et al., 2003). This suggests that K-turns at the terminal loop of box H/ACA sRNAs may be common in archaea. In the terminal loop structure, the 3 nucleotide asymmetric internal loop found in canonical K-turns is replaced by a loop of 4 to 7 nucleotides (Rozhdestvensky et al., 2003). Such terminal loops would be incapable of forming the Watson-Crick stem of a conventional K-turn and thus would lack the ability to form A-minor interactions. Thus, it appears that terminal loop L7Ae binding sites can fold into a K-turn-like structure without recourse to either the Watson-Crick stem or the A-minor interactions, two of the four defining features of K-turns.

In order to determine what allows sequences missing half of the canonical K-turn elements to form functional L7Ae binding sites, we have now solved the structure of *Methanococcus jannaschii* L7Ae protein bound to an RNA derived from a cognate site found in a terminal loop of an archaeal box H/ACA sRNA (Figure 1A). In our cocrystals, the stem loop oligonucleotide has formed a symmetric dimer in which two A•G paired stems flank a large internal loop. Two independent molecules of L7Ae bind to this dimeric RNA, each recognizing one A•G stem and one-half of the loop. We find that the RNA indeed folds into the characteristic structure with an extruded nucleotide that is flanked by the tandem A•G pairs, and is recognized by L7Ae in the same manner as conventional K-turns are. Thus, our 1.8 Å resolution structure demonstrates how the docking site for the phylogenetically widespread L7Ae family of proteins can be assembled from a minimal set of nucleotides. This minimal set may also define the structure of RNAs other than box H/ACA. In particular, eukaryotic box C/D snoRNAs and archaeal box C/D sRNAs contain a second binding site for L7Ae termed the box C′/D′ (Kiss-Laszlo et al., 1998). The majority of archaeal box C′/D′ elements appear to be terminal stem loops (snoRNA database, http://rna.wustl.edu/snoRNAdb/), and some have been shown to interact with L7Ae (Rashid et al., 2003; Tran et al., 2003). Furthermore, the binding of L7Ae to this type of site has been demonstrated to be essential for full activity of those RNPs (Tran et al., 2003).

Results and Discussion

Structure Determination

The central stem loop of the archaeal pseudouridylation guide sRNA afu-190 (Figure 1A) has previously been shown to bind to L7Ae with an apparent dissociation constant of ~100 nM (Rozhdestvensky et al., 2003). In order to obtain well-ordered cocrystals, oligonucleotides of various lengths that spanned the binding site were subjected to cocrystallization trials. Efficient in vitro binding of L7Ae to these oligonucleotides was confirmed by the electrophoretic mobility shift assay (data not shown). A 15 nucleotide RNA produced orthorhombic cocrystals that diffracted X-rays to 1.8 Å resolution.

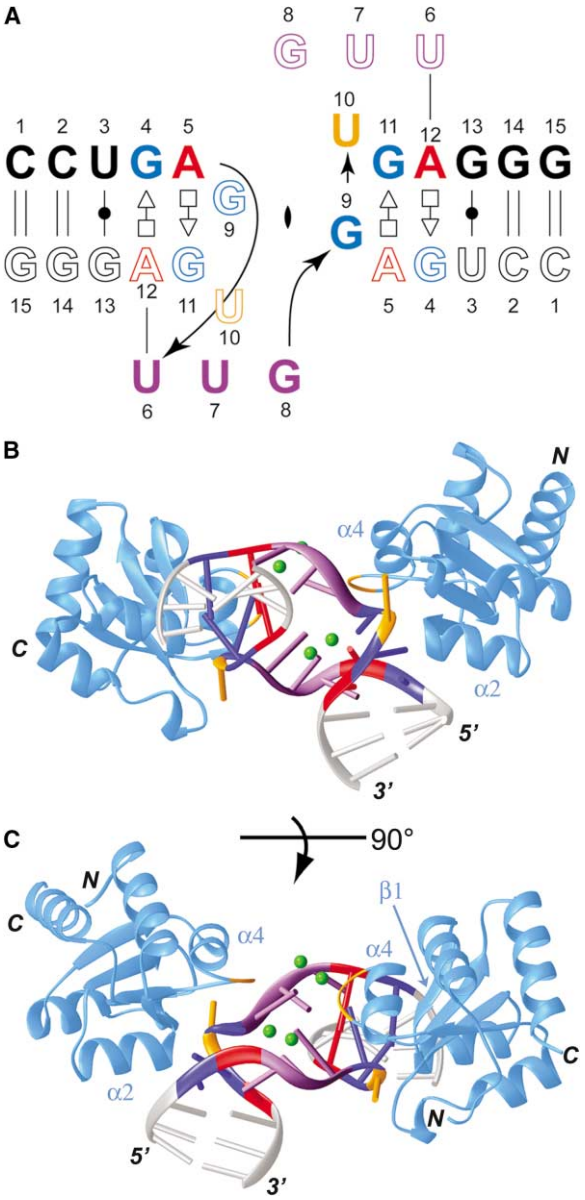


Figure 2. Overview of the L7Ae-RNA Complex

(A) Schematic secondary structure of the RNA in the cocrystal. The 2-fold symmetry axis of the duplex RNA is depicted by the lune. Each RNA strand is numbered from its 5′ end and depicted as either filled of open letters. Arrows indicate the path of one strand. The equivalent path of other strand has been omitted for clarity.

(B) Ribbon representation of L7Ae bound to the RNA (backbone shown as a ribbon, bases as cylinders). Four well-ordered calcium ions bound to the internal bulge are shown as green spheres. Color coding as in (A).

(C) View of the complex in the same representation as (B), but rotated 90° along the horizontal axis. The protein loop that follows helix 4, mentioned in the text, is colored yellow.

In order to solve the structure, uracil residues in the oligonucleotide were individually replaced with 5-bromouracil. One substitution, that at position 10 (numbering scheme in Figure 2A), produced crystals essentially isomorphous with the native form. The structure was solved by the multiwavelength anomalous dispersion (MAD) method. The crystallographic model has been refined to a free R factor of 22.5% and has an estimated precision of 0.20 Å (Experimental Procedures).

Overview of the Structure

In the cocrystals, instead of the terminal hairpin structure, the oligonucleotide has dimerized to produce an RNA with two A•G tandem stems (Figure 2A), both of which are bound by crystallographically independent L7Ae molecules that do not interact with each other (Figures 2B and 2C). The two L7Ae binding sites are separated by an internal loop that introduces a bend of ~130° into the RNA. Although noncrystallographic symmetry restraints were not used during refinement (because their use increased the free R factor), the three RNA and L7Ae molecules in the crystallographic asymmetric unit are very similar (root-mean-square differences [rmsd] of ~0.2 and ~0.4 Å, respectively).

Although our dimeric structure is not that of the authentic stem loop L7Ae binding site of the box H/ACA sRNA, we suggest that it nonetheless addresses the fundamental question of what constitutes the minimal K-turn for three reasons. First, the Watson-Crick stem of a canonical K-turn is absent from the duplex structure. Second, the two A•G stems of the RNA duplex do not interact with each other. Thus, the A-minor interaction of a canonical K-turn is absent. Third, the two RNA chains flanking the large internal loop separating the two A•G stems do not interact with each other. In this, they resemble a terminal loop, rather than the highly structured internal loop seen in canonical K-turns (Klein et al., 2001). The broader question of whether two terminal stem loop L7Ae binding sites in a box H/ACA sRNA, or two box H/ACA RNAs each with a single stem loop binding site, can dimerize in their biological context to form a structure like that present in our cocrystals, remains to be addressed.

As expected, the overall structure of L7Ae from *M. jannaschii* is very similar to previously reported structures of homologs. Superposition of *M. jannaschii* L7Ae with that from *Haloarcula marismortui* (Klein et al., 2001) produces an rmsd of 0.76 Å (for 115 Cα pairs). The rmsd between human 15.5 kDa (Vidovic et al., 2000) and *M. jannaschii* L7Ae is 1.0 Å (for 111 Cα pairs). The structure of *M. jannaschii* L7Ae is also very similar to the recently determined structure of L7Ae from *Archaeoglobus fulgidus* (rmsd of 0.67 Å for 112 Cα pairs) (H. Li, personal communication). The protein folds into a globular domain consisting of a central β sheet sandwiched between α helices. Five helices and four strands alternate with each other. RNA-protein interaction occurs primarily through strand 1, the N-terminal side of helix 2, and the loop after helix 4 (Figures 2B and 2C).

RNA Structure

The dimeric RNA in our cocrystals forms two independent A•G stems, each with the characteristic extruded U10 residue (Figure 2). The two A•G stems flank an internal loop or bulge that is spanned by two nucleotides, U7 and G8, which would form the loop in the native stem loop structure of the box H/ACA sRNA (Figure 1A). Consistent with the independent folding of each half of

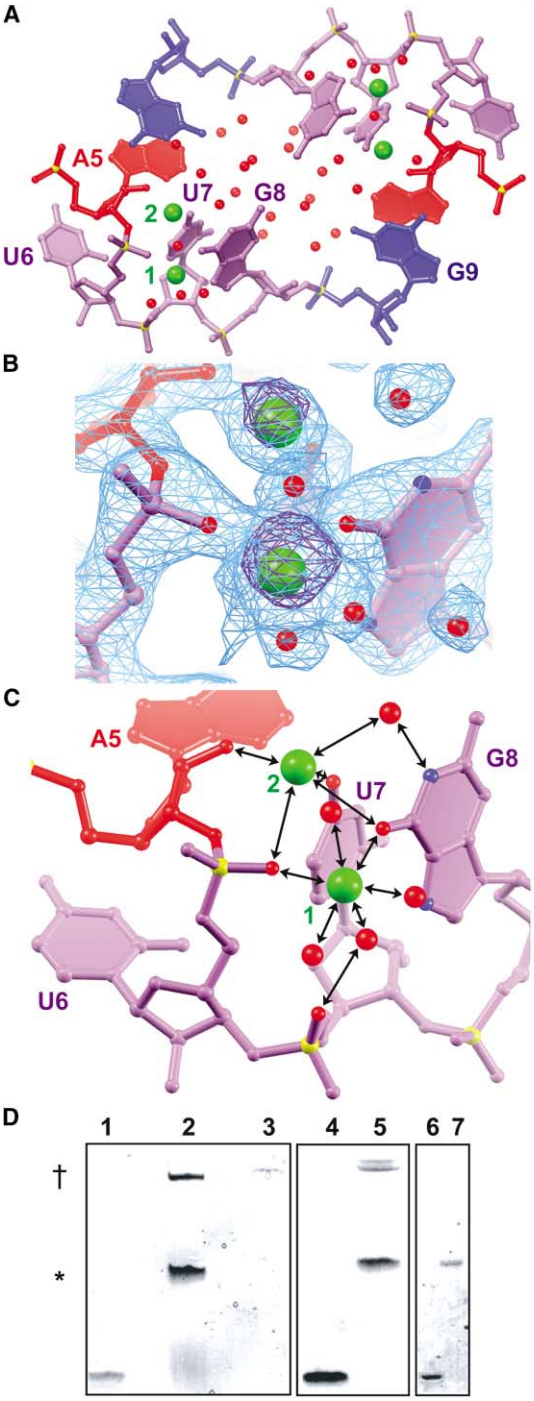


Figure 3. Structure of the Internal Loop

(A) The internal loop is filled with water molecules (red spheres) and coordinates four calcium ions (green spheres). The closest approach of G8 to its symmetry mate is 4 Å.

(B) Portion of the 1.8 Å resolution solvent-flattened MAD experimental electron density map contoured at 1.25 standard deviations above mean peak height (SD; blue mesh), superimposed on the refined crystallographic model. The magenta mesh shows the two highest peaks from a Fourier synthesis calculated with anomalous differences from crystal II and the MAD phases, contoured at 6.0 SD. The slight misalignment of the anomalous features and the calcium ions reflects differences in unit cell parameters of crystals II and III (Table 1). Note the clear experimental electron density for five of the six ligands of the lower calcium ion (labeled 1 in [A] and [C]).

the duplex RNA, U7 and G8 from one protomer do not interact with the symmetry-related nucleotides from the other. Instead, the loop nucleotides surround a water-filled cavity (Figure 3A). The nucleobases of U7 and G8 provide some of the ligands for two well-ordered calcium ions (Figure 3B). The two ions are separated by only 3.9 Å, and share two oxygen ligands. One of these is the carbonyl oxygen of G8 (Figure 3C).

Simultaneous occupancy of these two calcium ion sites is supported by two lines of evidence. First, the B factors for these two calcium ions are 32.7 $Å^2$ (SD 3.4 $Å^2$) and 32.6 $Å^2$ (SD 2.5 $Å^2$) (mean and standard deviation for the three protein-RNA complexes in the crystallographic asymmetric unit). These B factors are comparable to those of the ligands provided to the calcium ions by the RNA (28.0 $Å^2$ [SD 4.1 $Å^2$] for the phosphate oxygen of U6 and 29.7 $Å^2$ [SD 1.4 $Å^2$] for the carboxyl oxygen of nucleotide G8) (Figure 3C). Thus, the degree of order of the two calcium ions is comparable to each other and to that of the RNA ligands. Second, individual unrestrained occupancy refinement of the two calcium ions (six in the asymmetric unit), resulted in occupancies of ~1. Cations coordinated with similar geometry, in which two oxygen ligands are shared by two metal ions, have been reported previously for other high-resolution RNA structures. For instance, in the 1.5 Å resolution structure of the loop E dodecamer derived from 5S rRNA, Correll et al. found two Mg(II) ions that had this coordination geometry and were separated by only 2.7 Å (Correll et al., 1997).

Electrophoretic mobility shift analysis of the binding of our crystallization RNA to L7Ae, suggests that the RNA duplex observed in our cocrystal structure is stabilized by divalent metal ions (calcium or magnesium) but not by a monovalent ion (potassium; Figure 3D). A species that comigrates with the RNA-protein complex extracted from crystals forms in the presence of the former but not the latter. Although our crystals are grown in the presence of high concentrations of calcium ions,

(C) Coordination spheres of the two calcium ions. Six crystallographically ordered ligands surround calcium 1 with coordination distances (mean 2.3 Å, SD 0.06 Å) typical of octahedral calcium (Katz et al., 1996). Calcium ion 2 has only five ordered ligands. As the coordination distances are longer (2.9 Å), the crystallographic ligands could represent an average of multiple conformations. Note the two ligands (the phosphate oxygen of U6 and the carbonyl oxygen of G8) that are shared between the two ions. 3.9 Å separate the two calcium ions. Black double-headed arrows indicate coordinations or hydrogen bonds.

(D) Duplex complex formation in the presence of either divalent or monovalent cations assayed by the electrophoretic mobility shift. Lanes 1, 4, and 6 were loaded with the 15-mer RNA only. Lanes 2, 5, and 7 were loaded with a 1:1 mixture of the 15-mer RNA and the L7Ae protein. The protein RNA mixture was incubated with 2.5 mM calcium chloride (lane 2), 2.5 mM magnesium chloride (lane 5), or 7.5 mM potassium chloride (lane 7). In addition, the electrophoresis buffer contained the same cations at the same concentrations. Potassium chloride was added at 7.5 mM so that its ionic strength would be equivalent to that of 2.5 mM divalent cation. Lane 3 was loaded with dissolved cocrystals. Each square indicates a separate gel. In the absence of divalent, a stem-loop-L7Ae complex (1:1) is formed (*, lane 7). In contrast, in the presence of either Ca^{2+} or Mg^{2+}, a duplex complex (2:2) can be detected (†, lanes 2 and 5). The crystals contain the duplex complex exclusively (lane 3).

physiologic concentrations of magnesium ion (2.5 mM) were sufficient to form the duplex complex. Furthermore, the binding of L7Ae to a full-length box H/ACA sRNA (Figure 1A) was also stabilized by a divalent metal ion (2.5 mM magnesium chloride) by approximately 1.5-fold in the same assay (data not shown). Thus, the metal ion-bound RNA structure we observe in our crystals may form in a biological context.

Nucleotide U6 provides a ligand (a phosphate oxygen) that is shared between the two calcium ions (Figure 3C). The orientation of U6, with its nucleobase perpendicular to the bases of the loop nucleotides U7 and G8, is made possible by the formation of a base triple that drastically alters the path of the RNA backbone (Figures 2A and 4). In this triple, the base of U6 enters the minor groove of the A•G stem, forming a Watson-Crick pair with A12 (part of the A•G pair distal from the loop). Comparison of the our structure with the structure of *H. marismortui* L7Ae bound to rRNA (Figures 5A and 5B) shows that the ribosomal complex has a very similar base triple in the equivalent position (Klein et al., 2001). In that structure, the distal sheared pair is not the conventional A•G pair, but a G•U pair. The presence of the smaller pyrimidine base of U allows the intromission of an adenine into the major groove to complete a base triple. The trajectory of the RNA backbone is very similar in the *H. marismortui* structure and in our structure. This similarity extends to the perpendicular arrangement of the nucleotide equivalent to U7 (Figure 5B). In the ribosome, this nucleotide (G249) is the start of the Watson-Crick stem. Thus, it appears that the base triple may be enough to specify the trajectory of the Watson-Crick stem as it departs a canonical K-turn.

Preliminary analyses of L7Ae binding to terminal stem loop binding sites in a full-length box H/ACA sRNA suggest that base triple formation may not be critical for assembly of an L7Ae binding site in that context. Electrophoretic mobility shift assays using wild-type and mutant sRNAs incapable of forming the base triple (produced by mutagenizing U44 [Figure 1A] to either A or C) showed that the affinity of L7Ae for this binding site was decreased by only ~1.3-fold by the mutation (data not shown). Presumably, the very short length (three nucleotides) of the loop of the terminal binding site (residues 44–46 in Figure 1A; residue 47 is expected to stack on the A•G stem and residue 48 would be buried in the protein) results in low enough conformational entropy that further decreasing the loop length to two nucleotides as a result of base triple formation does not have a major energetic consequence.

The A•G stem of our structure conforms closely to those of canonical K-turns (Figure 4). The distal portion of the stem consists of Watson-Crick pairs. Adjacent and proximal to the loop is a wobble G•U pair, which is then followed by two tandem A•G pairs. On the 3′ side of the stem, U10 is extruded into solvent, and the nucleotide following this, G9, flips back onto the stem, to stack its nucleobase over the A of the sheared A5•G11 pair (Figure 4A). The two adenine bases of the sheared pairs participate in considerable cross-strand stacking, while the two guanine bases do not (Figure 4B). Instead, the latter bases hydrogen bond with the 2′-hydroxyl of each other's ribose moiety (Figure 4B).

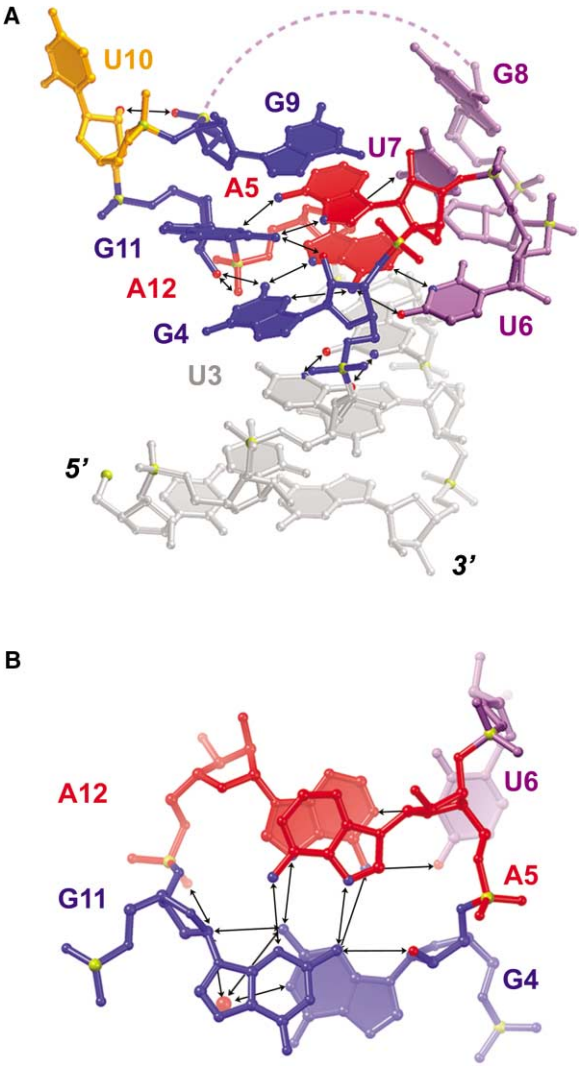


Figure 4. Structure of the L7Ae Binding Site

(A) Overview of the L7Ae binding site. Dotted line represents the trajectory of the RNA chain that would close off the terminal loop binding site (Figure 1A). L7Ae recognizes the tandem sheared A•G pairs from the major groove (left side) as well as the extruded U10 (yellow) in the loop.

(B) Structure of the tandem sheared A•G pairs viewed from the loop side (top in [A]). Note the substantial cross-strand stacking of the adenosines, and the base triple formed by the filled nucleotides (G4, A12, and U6).

RNA-Protein Interface

Association of one L7Ae molecule with RNA buries a total of 1500 Å^2 of solvent accessible surface area. L7Ae recognizes the RNA by interacting with two of its structural elements (Figure 6A), both of which have been shown to be protected from chemical modification by L7Ae binding to a terminal stem loop (Rozhdestvensky et al., 2003). First, the protein makes a number of contacts with the extruded loop nucleotide U10 (Figure 6B). Second, L7Ae recognizes the intricately folded structure of the A•G stem (stabilized by stacking interactions and the dense network of "tertiary" hydrogen bonds depicted with red arrows in Figure 6A) by docking against its

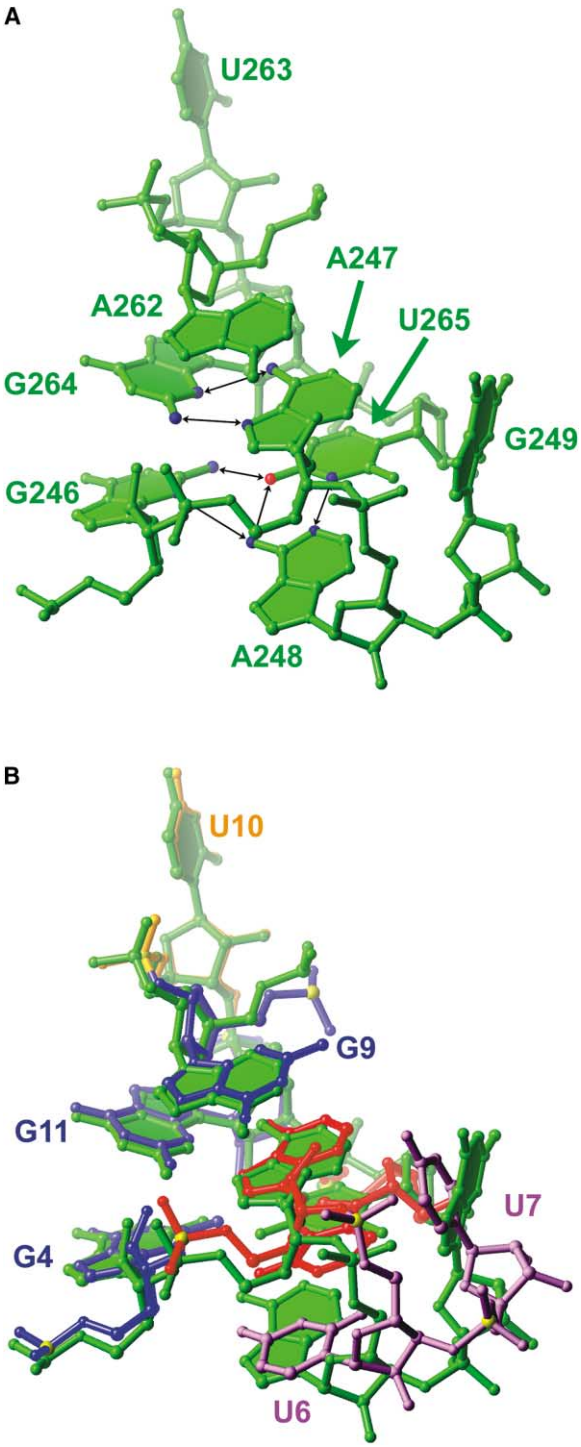


Figure 5. Comparison of the Structures of Base-Triple Containing K-Turns

(A) In the L7Ae binding site in the *H. marismortui* 50S ribosomal subunit (Klein et al., 2001), the distal A•G pair is replaced by the G246•U265 pair.

(B) Comparison of the *H. marismortui* and *M. jannaschii* L7Ae binding sites. The position of the two RNA chains is a result of superimposing the protein moieties of the complexes. Nucleotides numbered in this panel are from the *M. jannaschii* complex. The G4•A12•U6 base triple of the latter complex (also see Figure 4B) is nearly isosteric with the base triple of the ribosomal RNA (A). Note the similarity in the orientation of G249 from the rRNA and U7 from the box H/ACA sRNA sequence.

major groove (Figure 6C). The interaction of L7Ae with its cognate site has an interesting dichotomy, in that predominantly protein backbone atoms are used to recognize the base of U10 (Figure 6B), whereas protein side chains are responsible for the bulk of interactions with the A•G stem (Figure 6C).

Mutations of the extruded loop nucleotide corresponding to U10 to C have been shown to diminish *M. jannaschii* L7Ae binding to a box C/D sRNA by a factor of ~15 (Kuhn et al., 2002). The corresponding mutation to G abolished binding (Kuhn et al., 2002). Consistent with this, we observe direct hydrogen bonds to both the N3 imine and the O4 carbonyl of the nucleobase of U10 from L7Ae (Figure 6B). A cytosine at this position would lose the hydrogen bond that the O4 of U receives from the backbone amide of K53. Unless the pKa of the mutant C is perturbed, the hydrogen bond to backbone carbonyl of K53 would also be abolished. A purine at this position would clash sterically with the backbone of L7Ae. In addition, both phosphate oxygens of U10 are recognized by the protein. One of these was probed by modification into sulfur in a study of the eukaryotic homolog 15.5 kDa, and found to result in strong interference (Weinstein Szewczak et al., 2002).

Three amino acid residues of L7Ae appear to be primarily responsible for recognizing the characteristic structure of the A•G stem (Figures 6A and 6C). Glutamate 33 makes both main chain and side chain hydrogen bonds to the base of G11 (part of the A•G pair proximal to the loop). The carboxylate of this amino acid also hydrogen bonds to the ribose 2′-hydroxyl of G4 (part of the distal A•G pair). This simultaneous hydrogen bonding to the guanosines of both sheared A•G pairs uniquely specifies the geometry of the RNA recognized by E33. Asparagine 32 also makes hydrogen bonds that specify both guanosines. Its backbone amide donates a hydrogen bond to the base of G11, while its side chain hydrogen bonds to the base of G4. Finally, lysine 36 hydrogen bonds to the base of G4 and helps position the side chain of N32 through a water-mediated interaction. Since the two sheared guanines are fixed relative to each other by stacking and hydrogen bonding interactions (Figure 4B), and N32, E33, and K36 are constrained geometrically by being part of an α helix (Figure 6C), just three amino acids can provide a high degree of structure-specific recognition. There are additional interactions between the phosphate backbone of the RNA and L7Ae (for instance, both nonbridging phosphate oxygens of U3 are recognized by K36 and R40, Figure 6A), but overall, the footprint of the protein on the RNA is restricted, and the major groove of the A•G stem is quite heavily hydrated.

Not surprising, mutations of the nucleotides comprising the sheared A•G pairs result in complete loss of protein binding in vitro (Rozhdestvensky et al., 2003; Kuhn et al., 2002; Rashid et al., 2003) and loss-of-function in vivo (Marmier-Gourrier et al., 2003). Modification analysis of 15.5 kDa interaction with a box C/D snoRNA both in vitro and in vivo revealed that alteration of the pro-Rp oxygen of G4, the O6 of A5, the N2 of G11, and the O6 of A12 (our nucleotide numbering scheme) interfered with the function of the RNA (Weinstein Szewczak et al., 2002). The replacement of the nucleobase

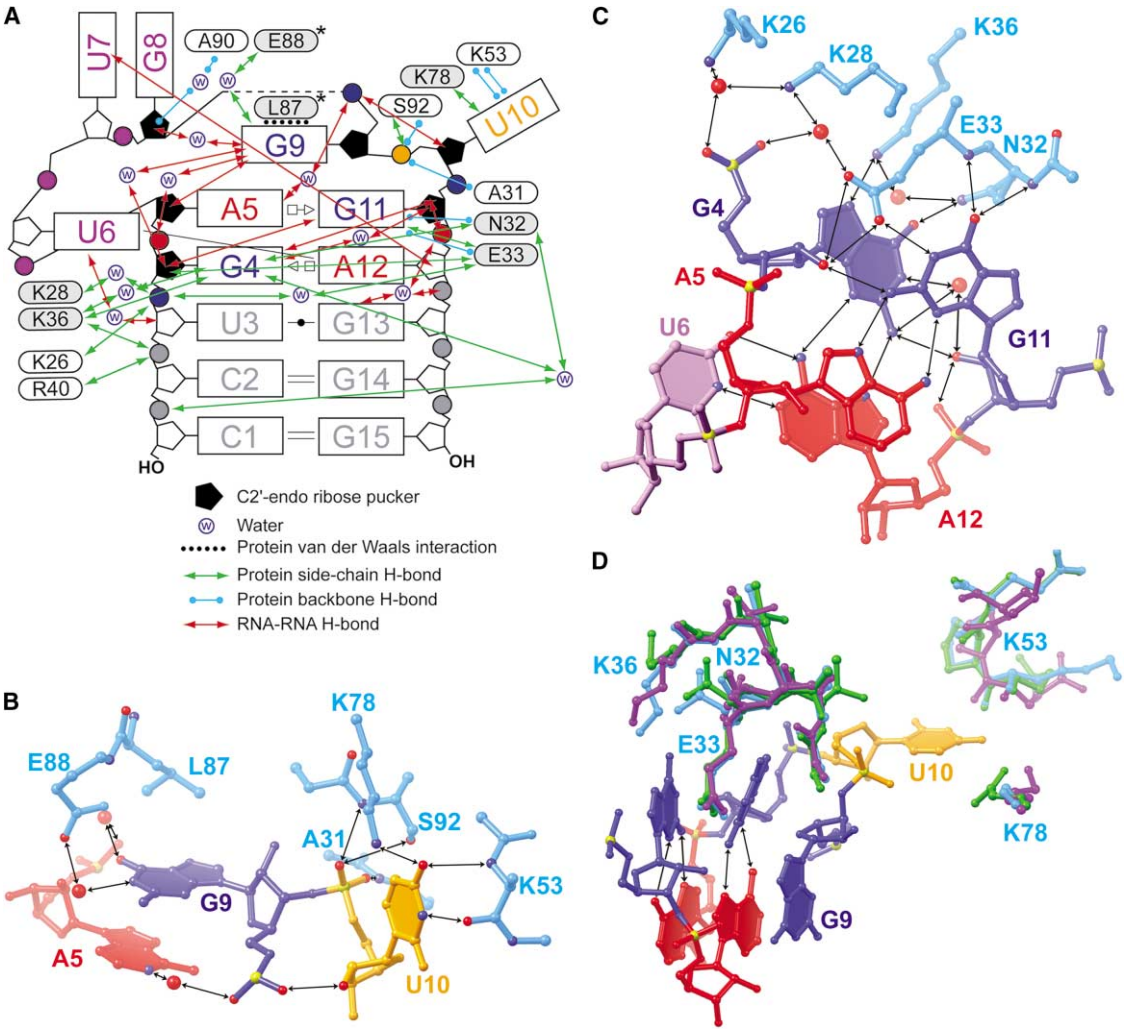


Figure 6. Protein-RNA Interface

(A) Summary of protein-RNA and RNA-RNA interactions. Amino acid residues that are conserved between archaea and eukaryotes are shown as gray-filled oval. The asterisks indicate conservation of an amino acid residue only within archaea.

(B) L7Ae recognition of the extruded U10.

(C) Protein recognition of the A•G stem. The protein recognizes the major groove faces of the sheared G primarily through interaction with the conserved residues N32, E33 and K36.

(D) Comparison of the RNA complexes of *M. jannaschii* L7Ae (light blue, this work), *H. marismortui* L7Ae (Klein et al., 2001) (green), and human 15.5 kDa (Vidovic et al., 2000) (purple). Portions of the RNA from the *M. jannaschii* complex are shown for reference.

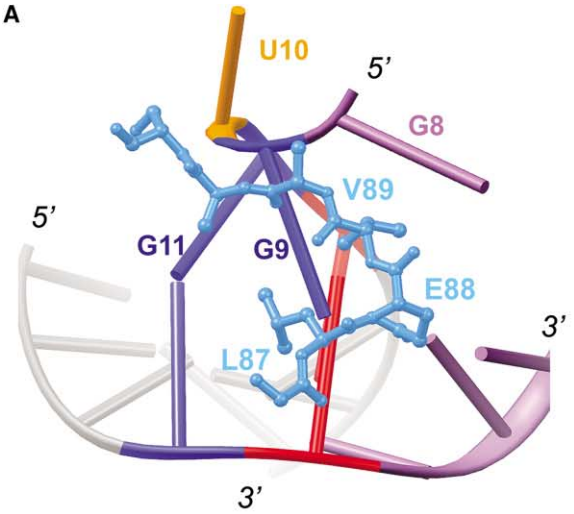
atoms would compromise the sheared A•G pairs while the replacement of the phosphate oxygen with sulfur would affect its interaction, through water, with the side chains of K28 and E33 (Figure 6C).

Comparison of the *M. jannaschii* L7Ae complex with the *H. marismortui* L7Ae in complex with K-turn 15 of the 23S RNA (Klein et al., 2001) and with the human 15.5 kDa protein bound to a cognate site on U4 snRNA (Vidovic et al., 2000) shows that the general features used for recognition by all three proteins are highly conserved (Figure 6D). Recognition of the tandem sheared A•G pairs employs residues equivalent to N32, E33, and K36 in exactly the same positions in helix 2 in all three proteins. Recognition of the extruded U10 relies on the same set of protein backbone atoms in all three cases. A structure of the *A. fulgidus* L7Ae bound to an archaeal box C/D sRNA shows an identical recognition mechanism (H. Li, personal communication).

RNA Recognition by L7Ae and Its Eukaryotic Homologs

Recently, it was demonstrated that whereas archaeal L7Ae can bind to both conventional K-turns with a Watson-Crick stem and to stem loop K-turns lacking that stem, the homologous eukaryotic protein 15.5 kDa could only bind to the former (Tran et al., 2003). Comparison of the L7Ae complex with the 15.5 kDa complex suggests a molecular basis for this difference. The archaeal L7Ae protein interacts with the nucleobase of G9 by making van der Waals contacts with L87 and a water-mediated

A



B

Archaeoglobus fulgidus (87)	...GIEVPCA...
Haloarcula marismortui (89)	...GLEVGSA...
Methanococcus jannaschii (86)	...GLEVAAS...
Human 15.5kD (94)	...GVSRPVI...
Mouse 15.5kD (94)	...GVSRPVI...
Saccharomyces cerevisiae Snu13p (92)	...GVSRPVI...
Arabidopsis thaliana L7Ae-like (94)	...GVTRPVI...
Cryptosporidium parvum (100)	...GVSRPVI...
Human Nhp2 (119)	...GSKRPTC...
Saccharomyces cerevisiae Nhp2 (131)	...ATKRPTS...

C

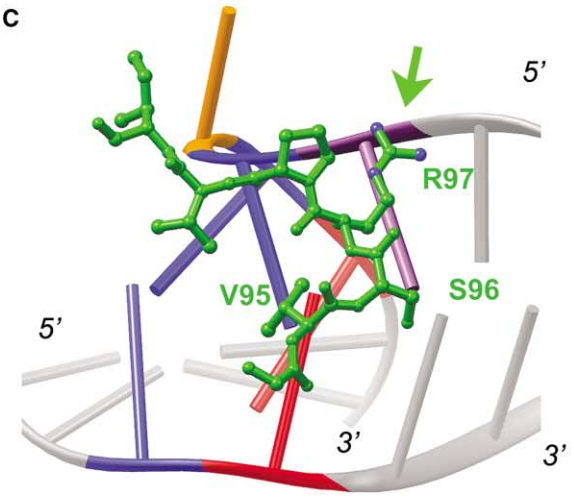


Figure 7. Molecular Basis for the Different RNA Structure Preferences of L7Ae and the Eukaryotic Homolog 15.5 kDa

(A) Detail of the *M. jannaschii* L7Ae-RNA complex, illustrating the contrasting disposition of the side chains of the LEV tripeptide. The RNA in (C) is depicted so that nucleotides that are structurally equivalent to those in (A) have the same color.

(B) Alignment of the sequences of the loop that follows helix 4 of representative archaeal and eukaryotic sequences. Note the high conservation of the tripeptide LEV in the former and its replacement with VSR in the latter (highlighted in gray). In parentheses are the numbers of the first amino acid residue shown in the alignment for each protein.

(C) Detail of the VSR tripeptide and its interaction with RNA in the structure of the 15.5 kDa RNA complex (Vidovic et al., 2000). The long side chain of the arginine of the VSR tripeptide crosses the loop to contact the last phosphate of the Watson-Crick stem (arrow).

interaction through the side chain of E88 (Figures 6A and 6B). The next amino acid in the protein chain, V89, is tucked back into the protein and forms one of the sides of the pocket that receives the extruded base of U10 (Figure 7A). All three amino acids, L, E, and V, are highly conserved among the archaeal L7Ae proteins (Figure 7B), and participate in equivalent interactions in the known structures.

In the eukaryotic proteins, in contrast, this LEV tripeptide (which forms the loop following helix 4, in yellow in Figure 2C) is replaced by another sequence, VSR, that is absolutely conserved from *Cryptosporidium* to human (Figure 7B). The structure of the human 15.5 kDa protein in complex with RNA (Vidovic et al., 2000) shows that the valine of the eukaryotic tripeptide interacts with the base equivalent to G9 much as the archaeal L87 does (Figures 7A and 7C). However, the arginine of the third position of the eukaryotic sequence makes an interaction that is absent in the archaeal complexes (L7Ae has a valine in this position). The long arginine side chain crosses the loop of the K-turn and makes a salt bridge with a phosphate at the proximal end of the Watson-Crick stem (green arrow, Figure 7C). Thus, to a first approximation, the RNA binding affinity of the eukaryotic proteins may be modulated not only by the presence of the Watson-Crick stem but also by the exact angle at which this stem departs from the loop of the K-turn. In contrast, archaeal L7Ae proteins do not make this—or any other—interactions with the Watson-Crick stem and thus can bind to both canonical and stem loop type K-turns (the latter lacking the Watson-Crick stem). Our structural comparison suggests that the arginine of the conserved eukaryotic loop sequence contributes significantly to the increased specificity (Tran et al., 2003) of the eukaryotic K-turn binding proteins.

The Minimal K-Turn

Determination of the high-resolution structure of L7Ae in complex with a minimal binding site has allowed us to dissect the K-turn into its constituents. This structure demonstrates that the A•G stem and the extruded U that characterize the binding sites of L7Ae can be arranged into the correct conformation for recognition without the presence of a conventional Watson-Crick stem or of A-minor interactions between the A•G and the Watson-Crick stems. Further structural analyses will be needed to establish whether minimal K-turns exist in RNAs other than box H/ACA sRNAs with terminal L7Ae binding sites. In particular, it is likely that binding of L7Ae to box C′/D′ elements of methylation guide RNAs (Rashid et al., 2003; Tran et al., 2003) results in the RNA adopting the same conformation that is present in our cocrystals.

Experimental Procedures

Protein and RNA Preparation

Plasmid pETmjL7Ae is a pET16b (Novagen) derivative that encodes the L7Ae protein from *Methanococcus jannaschii* (residues 1–117).

Table 1. Crystallographic Statistics

Diffraction Data					
Data Set	Crystal I			Crystal II	Crystal III
Resolution range (Å)	50.0–2.50	50.0–2.33	50.0–2.48	99.0–2.20	30.0–1.80
Last shell (Å)	2.59–2.50	2.41–2.33	2.57–2.48	2.28–2.20	1.86–1.80
Unit cell					
a (Å)	106.01			106.01	106.97
b (Å)	140.17			140.59	141.30
c (Å)	163.98			164.08	162.87
	λ1	λ2	λ3		
Wavelength (Å)	0.9201	0.9199	0.9150	1.5418	1.0000
Reflections observed	88,150	102,712	178,600	352,841	386,751
Unique	21,249	23,463	21,535	31,507	56,829
Completeness (%)	99.4	88.4	99.5	99.8	99.7
(last shell)	(99.8)	(89.4)	(99.9)	(99.7)	(100.0)
$<I>/<\sigma(I)>$ (last shell)	33.2 (8.4)	32.1 (7.4)	39.7 (9.0)	34.5 (6.5)	31.2 (3.2)
R_{sym}[a] (%) (last shell)	3.6 (12.9)	4.0 (15.5)	5.1 (18.3)	9.7 (45.1)	6.0 (50.1)
MAD Analysis					
Mean figure of merit[b]	Resolution range (Å)				
0.48	30–2.5				
0.23	2.58–2.5				

Numbers in parentheses refer to the last shell.

[a] $R_{sym} = \Sigma|I - <I>|/\Sigma I$, where I is the observed intensity and $<I>$ is the statistically weighted absolute intensity of multiple measurements of symmetry related reflections.

[b] Mean figure of merit $= <\Sigma P(\alpha)e^{i\alpha}/\Sigma P(\alpha)>$, where α is the phase and $P(\alpha)$ is the phase-probability distribution.

Unlike an expression construct that was previously reported (Kuhn et al., 2002), our construct does not encode for any affinity tags. The coding sequence was generated by PCR amplification from genomic DNA (purchased from ATTC). L7Ae was expressed in *Escherichia coli* BL21-CodonPlus(DE3)-RIL (Stratagene). The lysate was heated at 65°C for 20 min in a buffer containing 400 mM KCl, 20 mM Tris HCl (pH 8.0), 0.5 mM PMSF, 0.5% aprotinin, 5 μg/ml leupeptin, and 5% glycerol. The supernatant was fractionated with polyethyleneimine and ammonium sulfate (Burgess, 1991). L7Ae was dissolved in buffer A (1.5 M ammonium sulfate, 20 mM HEPES-KOH (pH 7.5), 2 mM DTT, 0.5 mM EDTA) and loaded onto a hydrophobic interaction column (Phenyl Sepharose 6 Fast Flow High Sub, Amersham) equilibrated in buffer A. The protein was eluted with a linear gradient to buffer B (20 mM HEPES-KOH (pH 7.5), 2 mM DTT, 0.5 mM EDTA). L7Ae was further purified by size-exclusion chromatography (Superdex 75, Amersham) using buffer C (150 mM KCl, 20 mM HEPES-KOH (pH 7.5), 2% glycerol, 2 mM DTT, 0.5 mM EDTA). The protein was concentrated and stored at −80°C in buffer C. Coomassie staining of serial dilutions of L7Ae analyzed on SDS polyacrylamide gels indicate that our L7Ae preparation is at least 99.6% pure. Mass spectrometry (MALDI-TOF) confirmed the molecular mass of L7Ae, missing the initiation methionine but otherwise intact. RNA oligonucleotides were purchased from Dharmacon Research, deprotected according to the manufacturer's instructions, and used without further purification.

Crystallization and Data Collection

L7Ae was mixed with RNA in a 1:1.5 protein:RNA molar ratio at a final complex concentration of 0.5 mM in buffer C, and incubated at 22°C for 1 hr. Crystals were grown at 22°C by the hanging-drop vapor diffusion method by mixing 1.5 μl each of complex and a reservoir solution consisting of 50–100 mM $CaCl_2$, 100 mM HEPES-KOH (pH 7.5), 150 mM KCl, and 16%–18% (v/v) PEG400 (or 16%–18% PEG550). Cocrystals (containing either unmodified RNA or 5-bromouracil containing RNA), which grew within a week to typical dimensions of 200 × 150 × 40 μm^3, were transferred to a solution containing 30% PEG 400 or PEG 550 (but having otherwise the same composition as the reservoir), and flash cooled in liquid nitrogen. The cocrystals (space group F222, cell parameters in Table 1) contain three protein-RNA complexes per asymmetric unit. Crystal I contained RNA with 5-bromouracil at position 10. Crystals II and III contained unmodified RNA. Data from crystals I and III were collected at 100 K at beamline 5.0.2 of the Advanced Light Source (ALS, Lawrence Berkeley National Laboratory). Data from crystal II were collected using the inverse-beam method in the home laboratory with X-radiation from a rotating anode equipped with multilayer optics. Diffraction data were reduced with the HKL package (Otwinowski and Minor, 1997).

Phase Determination and Structure Refinement

Three bromine sites were located, their parameters refined, and phases calculated using three-wavelength data from crystal I (Table 1) with SOLVE (Terwilliger and Berendzen, 1999). Solvent flipping, histogram matching, and phase extension to 1.8 Å resolution (crystal III data) with CNS (Brünger et al., 1998) produced an electron density map (Figure 3B) into which all the RNA, most of the protein residues, and the bulk of ordered ions and water molecules could be built unambiguously with program O (Jones et al., 1991). Assignment of density features to calcium ions was confirmed by inspection of anomalous Fourier syntheses calculated with amplitudes from crystal II and MAD phases. Rounds of energy minimization, simulated annealing, and restrained individual B factor refinement using a maximum likelihood target, interspersed with manual rebuilding, produced the current model (R_{free} = 22.5%, R_{work} = 20.7%.) Refinement was against all crystal III data ($|F| > 0$) and all experimental MAD phase probability distributions (5,778 and 51,018 structure factors in the test and work sets, respectively). A solvent mask and an overall anisotropic B factor correction were used throughout. The crystallographic model comprises all nucleotides in the three RNA chains in the asymmetric unit, amino acid residues 2–116, 7–115, and 5–116 of protein chains A, B, and C, respectively, 272 water molecules, 10 calcium ions, and 1 potassium ion (3,698 nonhydrogen atoms). The cross-validated σA mean coordinate error is 0.20 Å. The mean B factors for RNA, protein, water, and cations are 25.7, 35.2, 37.5, and 37.5 $Å^2$, respectively. The model has rmsd of 0.004 Å and 1.16° from ideal bond lengths and angles, respectively. The mean real-space R factors (in a σA-weighted composite simulated-annealing omit $2|F_o| - |F_c|$ map) are 9.23% and 3.23% for protein and RNA residues, respectively. 96.7% of the amino acid residues lie in the most favored regions of the Ramachandran plot, with the remaining residues in additional allowed regions. There are no amino acid residues with disallowed backbone conformations. Except for line drawings, all figures were

prepared with RIBBONS (Carson, 1991). Shown in the figures are protein chains A and B, and RNA chains D and E. Figures 6 and 7 are representations of protein chain A, which is the best ordered of the three polypeptides in the asymmetric unit.

Electrophoretic Mobility Shift Assays

Complexes were prepared by mixing the 15 nucleotide RNA used for cocrystallization and *M. jannaschii* L7Ae at a final concentration of 0.1 mM in a solution containing 20 mM HEPES-KOH (pH 7.5) supplemented with either 2.5 mM calcium chloride, 2.5 mM magnesium chloride, or 7.5 mM potassium chloride and incubated at 22°C for 1 hr. Before mixing with the protein, the RNA was heated for 2 min at 75°C in water, and immediately cooled in ice water for 5 min. The complexes were separated from free RNA on 15% native PAGE gels run at 26°C in a buffer (pH 7.5) containing 16.5 mM Tris, 33 mM HEPES, and 0.05 mM EDTA also supplemented with the same concentrations of ions used in the binding buffer. Gels were fixed and stained in 40% (v/v) methanol, 1% (v/v) acetic acid, and 0.1% (w/v) Toluidine blue O (Sigma) for 30 min; then, destained in 40% methanol and 1% acetic acid for 2 hr.

Acknowledgments

We thank the staff of ALS beamline 5.0.2 for data collection support, J. Bolduc and B. Shen for crystallographic support, B. Lunde for assistance during model building and refinement, and T. Edwards, C. Hoang, D. Kline, J. Pitt, P. Rupert, C. Spiegel, and H. Xiao for discussions. We thank Dr. Hong Li (Florida State University) for sharing atomic coordinates prior to publication. T.H. is a trainee of the Viral Oncology Training Grant from the National Institutes of Health (NIH) to the University of Washington (CA09229). In addition, this work was supported by grants from the NIH (GM63576 and RR15943) and the Rita Allen Foundation to A.R.F. Access to ALS beamline 5.0.2 as part of the principal research consortium was made possible by general support from the Fred Hutchinson Cancer Research Center. A.R.F. is a Rita Allen Foundation Scholar and a W.M. Keck Foundation Distinguished Young Scholar in Medical Research.

Received: October 16, 2003
Revised: December 17, 2003
Accepted: February 11, 2004
Published: May 11, 2004

References

Bachellerie, J.P., Cavaille, J., and Huttenhofer, A. (2002). The expanding snoRNA world. Biochimie *84*, 775–790.

Badis, G., Fromont-Racine, M., and Jacquier, A. (2003). A snoRNA that guides the two most conserved pseudouridine modifications within rRNA confers a growth advantage in yeast. RNA *9*, 771–779.

Bousquet-Antonelli, C., Henry, Y., Gélugne J,P., Caizergues-Ferrer, M., and Kiss, T. (1997). A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. EMBO J. *16*, 4770–4776.

Brünger, A.T., Adams, P.D., Clore, G.M., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., et al. (1998). Crystallography and NMR system: a new software system for macromolecular structure determination. Acta Crystallogr. D *54*, 905–921.

Burgess, R.R. (1991). Use of polyethyleneimine in purification of DNA-binding proteins. Methods Enzymol. *208*, 3–10.

Carson, M. (1991). Ribbons 2.0. J. Appl. Crystallogr. *24*, 958–961.

Correll, C.C., Freeborn, B., Moore, P.B., and Steitz, T.A. (1997). Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. Cell *91*, 705–712.

Galardi, S., Fatica, A., Bachi, A., Scaloni, A., Presutti, C., and Bozzoni, I. (2002). Purified box C/D snoRNPs are able to reproduce site-specific 2′-O-methylation of target RNA in vitro. Mol. Cell. Biol. *22*, 6663–6668.

Ganot, P., Bortolin, M.-L., and Kiss, T. (1997). Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. Cell *89*, 799–809.

Gaspin, C., Cavaille, J., Erauso, G., and Bachellerie, J.P. (2000). Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the Pyrococcus genomes. J. Mol. Biol. *297*, 895–906.

Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gelugne, J.P., and Caizergues-Ferrer, M. (1998). Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. EMBO J. *17*, 7078–7090.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Katz, A.K., Glusker, J.P., Beebe, S.A., and Bock, C.W. (1996). Calcium ion coordination: a comparison with that of beryllium, magnesium, and zinc. J. Am. Chem. Soc. *118*, 5752–5763.

King, T.H., Liu, B., McCully, R.R., and Fournier, M.J. (2003). Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. Mol. Cell *11*, 425–435.

Kiss-Laszlo, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell *85*, 1077–1088.

Kiss-Laszlo, Z., Henry, Y., and Kiss, T. (1998). Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. EMBO J. *17*, 797–807.

Klein, D.J., Schmeing, T.M., Moore, P.B., and Steitz, T.A. (2001). The kink-turn: a new RNA secondary structure motif. EMBO J. *20*, 4214–4221.

Klein, R.J., Misulovin, Z., and Eddy, S.R. (2002). Noncoding RNA genes identified in AT-rich hyperthermophiles. Proc. Natl. Acad. Sci. USA *99*, 7542–7547.

Koonin, E.V., Bork, P., and Sander, C. (1994). A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins, and a ribosome modification enzyme. Nucleic Acids Res. *22*, 2166–2167.

Kuhn, J.F., Tran, E.J., and Maxwell, E.S. (2002). Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein. Nucleic Acids Res. *30*, 931–941.

Leontis, N.B., and Westhof, E. (2001). Geometric nomenclature and classification of RNA base pairs. RNA *7*, 499–512.

Mao, H., White, S.A., and Williamson, J.R. (1999). A novel loop-loop recognition motif in the yeast ribosomal protein L30 autoregulatory RNA complex. Nat. Struct. Biol. *6*, 1139–1147.

Marmier-Gourrier, N., Clery, A., Senty-Segault, V., Charpentier, B., Schlotter, F., Leclerc, F., Fournier, R., and Branlant, C. (2003). A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. RNA *9*, 821–838.

Ni, J., Tien, A.L., and Fournier, M.J. (1997). Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. Cell *89*, 565–573.

Nicoloso, M., Qu, L.H., Michot, B., and Bachellerie, J.P. (1996). Intron-encoded, antisense small nucleolar RNAs: the characterization of nine novel species points to their direct role as guides for the 2′-O-ribose methylation of rRNAs. J. Mol. Biol. *260*, 178–195.

Nissen, P., Ippolito, J.A., Ban, N., Moore, P.B., and Steitz, T.A. (2001). RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. Proc. Natl. Acad. Sci. USA *98*, 4899–4903.

Omer, A.D., Lowe, T.M., Russell, A.G., Ebhardt, H., Eddy, S.R., and Dennis, P.P. (2000). Homologs of small nucleolar RNAs in archaea. Science *288*, 517–522.

Omer, A.D., Ziesche, S., Ebhardt, H., and Dennis, P.P. (2002). In vitro reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex. Proc. Natl. Acad. Sci. USA *99*, 5289–5294.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Rashid, R., Aittaleb, M., Chen, Q., Spiegel, K., Demeler, B., and Li, H. (2003). Functional requirement for symmetric assembly of archaeal box C/D small ribonucleoprotein particles. J. Mol. Biol. *333*, 295–306.

Rozhdestvensky, T.S., Tang, T.H., Tchirkova, I.V., Brosius, J., Bachellerie, J.P., and Huttenhofer, A. (2003). Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. Nucleic Acids Res. *31*, 869–877.

Tang, T.H., Bachellerie, J.P., Rozhdestvensky, T., Bortolin, M.L., Huber, H., Drungowski, M., Elge, T., Brosius, J., and Huttenhofer, A. (2002). Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus. Proc. Natl. Acad. Sci. USA *99*, 7536–7541.

Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr. D *55*, 849–861.

Tran, E.J., Zhang, X., and Maxwell, E.S. (2003). Efficient RNA 2′-O-methylation requires juxtaposed and symmetrically assembled archaeal box C/D and C′/D′ RNPs. EMBO J. *22*, 3930–3940.

Tycowski, K.T., Smith, C.M., Shu, M.D., and Steitz, J.A. (1996). A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in Xenopus. Proc. Natl. Acad. Sci. USA *93*, 14480–14485.

Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R., and Ficner, R. (2000). Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. Mol. Cell *6*, 1331–1342.

Watanabe, Y., and Gray, M.W. (2000). Evolutionary appearance of genes encoding proteins associated with box H/ACA snoRNAs: Cbf5p in Euglena gracilis, an early diverging eukaryote, and candidate Gar1p and Nop10p homologs in archaebacteria. Nucleic Acids Res. *28*, 2342–2352.

Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M., and Luhrmann, R. (1998). Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. RNA *4*, 1549–1568.

Watkins, N.J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., and Luhrmann, R. (2000). A common core RNP structure shared between the small nucleoar box C/D RNPs and the spliceosomal U4 snRNP. Cell *103*, 457–466.

Weinstein Szewczak, L.B., DeGregorio, S.J., Strobel, S.A., and Steitz, J.A. (2002). Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP. Chem. Biol. *9*, 1095–1107.

Winkler, W.C., Grundy, F.J., Murphy, B.A., and Henkin, T.M. (2001). The GA motif: an RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs. RNA *7*, 1165–1172.

Accession Numbers

Atomic coordinates and structure factor amplitudes have been deposited with the Protein Data Bank (accession code 1SDS).